

ACACCATTG TCTTCATGTA ACCCCATTAG CTATACCCTC TAGTGCAAGG AAACCATAGG
 10 20 30 40 50 60
 GCCTAGGTCA CACCATGAGG CTGCNCTTAC AAGTTATGCA AAAACTATGG ACTTGGGAGA
 70 80 90 100 110 120
 CCTGTGCGTA ACAACATCAC ACNCCAAATT TAACCAGCTC TCCCCATAAC AGCACGCTCA
 130 140 150 160 170 180
 TGTGTTACTG AGGAAATGCC TGTGGATTGG AGTGTGTTCT GTGTGCAGGA GGCTGGTCCA
 190 200 210 220 230 240
 GGTTTCACTT CTGCAGGACA CTGGACGTTT CCCAAAACCA GCAGACTTTC CCCACGTGCA
 250 260 270 280 290 300
 CACACACCCC TTCTCATTTT GCCTCTACAT CCATATCCAC TGGGCCCTTC AGGCACCTAC
 310 320 330 340 350 360
 TAATGCCCTA GAACCTAAAA CCATCATCTG GGGCCCAGTT CCCTGAATGG CCCTAATCTC
 370 380 390 400 410 420
 TTCCTCTGCT GGAATGAGTC CAGTGCCAC TTCCTCCAAC GGTGAAATTG CTGGGCTGCT
 430 440 450 460 470 480
 ACAGATCAGG AACTCACTGC TTCCTCATAG GGCAGCCGA CTTCCTGCT CTGCAACAGC
 490 500 510 520 530 540
 GACCACCCCT AGCGAGGCTT GAGATGCCTC TTGCCTCCTT AAGACTGAGG GAGACGCTTC
 550 560 570 580 590 600
 AGCTCTCACT CCACTGCCCC AAGTCCTCCA CAGCGCGGTG CCTGCTGCCT TCACACAGAG
 610 620 630 640 650 660
 CTGCAGGGGN AGGTCCTGTG TATCCGGCCT GCTGGACCAG CGCTGTGCAC AACCTCCCA
 670 680 690 700 710 720
 TGGCAACAGT GGCTGCCCCG CCTGCACACT GGGCTTGGCA ACCTCGCTGT AGGTATTTAT
 730 740 750 760 770 780
 TCCCTCAGGA GTGACTGCAT TCTTTTCCCA TTTCCAGAAA ACTGATGCCA TTTACCTCAC
 790 800 810 820 830 840
 TATGAGGAGG AGGAGGAGGA GGAGGGTGGA GAGTGGTACA TTTTAAATG TGCATATTCT
 850 860 870 880 890 900
 TCCCTAGGAC TCCCCCTCAA ATAACCCAGG AGGGACCATA CCAGCTCATT CCTGTGTATC
 910 920 930 940 950 960
 CCAAGCATAN GAGTAATCAT CCCACTCATG CTGAGTGTAT GGTGGCCATT AAGCCTGCC
 970 980 990 1000 1010 1020

Figure 1

TGAAGCTGGCT TTAGAACAAG GTGTTTGAGC ACACAGCACC GTCTTGCTGC CACCTTGGCC
 1030 1040 1050 1060 1070 1080

CCCTCCCTTG TGAGACCTCT GAGACACATT NAGGTCTCAC CTAAAAATCT CAGGATTTCT
 1090 1100 1110 1120 1130 1140

AGGCCCAAAN CGGTCCTAAA AAATTGTCA GTCTGAACTC TCTAAGGTCA AGAGAAGAGG
 1150 1160 1170 1180 1190 1200

TGGTTGCTCC CTCTAAGAAA CCACATGTTG CATGTACATC CTTAATTCCG GAAAGTCCAA
 1210 1220 1230 1240 1250 1260

CAAACCTGCC CTGCTTAGCA ACACAAGCCG AGGTGGTACT CCTCTCACCC GGGCATTCTC
 1270 1280 1290 1300 1310 1320

CAACACACCT GTTTGTCCAA ACAGCTTTGA TTTGTTTTTA TAGTTGGACC CCAGGTTCCC
 1330 1340 1350 1360 1370 1380

AGGAGGCTGG TTCAGGCCAT ATTCCAAATC CTCATCTGTG TGTGAGTGGC ATTCTTAGCC
 1390 1400 1410 1420 1430 1440

TAGCCTCCTT ACAGGGTGA TACTATGATA CACAGCCAGG CTGTCCCAGT GGCTTTCAAT
 1450 1460 1470 1480 1490 1500

ATTCTTTTGG TCCAGATAGT TCAGCCTCAG CACCAAGTGA GGCATCACAG GGTCAATTGT
 1510 1520 1530 1540 1550 1560

CTTAGGAGTC ATGGAGAATT CATAGTTGGT AGCTACCTGG GCCTGGCCAG GGCTGACCAT
 1570 1580 1590 1600 1610 1620

AGACAAGGCA TCCCTCTGTG AACTCCTATT TTAATGCCAG CTTCCCAACA AATTTCTCAA
 1630 1640 1650 1660 1670 1680

CAAT box

CTGCTCTTAC CAGCAGGTAT TTAAACTACT CAATAGAAAG TAACCCTGAA AATTAGGACA
 1690 1700 1710 1720 1730 1740

TATA box

CCTGTTCCTA AAAGACCCTT AAATAGGGGA AGTCCTTTCTN CTGCTTGTGC ACAGCTGCTG
 1750 1760 1770 1780 1790 1800

| ->mRNA -----

ATGTGGCAAC ATGAGGCCTG GGACAGGGGA CTGTCCTCTG CCCACTCTGG TAGCCTCAGC
 1810 1820 1830 1840 1850 1860

Spsite
 -- exon 1 ---->#####

TAGCTTAACA ATCTGTCAGT AATACAATAC AAAACTTAAA CTTTCATACT GCGGTTCCAC
 1870 1880 1890 1900 1910 1920

CCAGGAAGCT GTGTTCCCAA TCTGACCCGT GATTATGGGG CCACCTCAGA GGGNACCCAG
 1930 1940 1950 1960 1970 1980

Figure 1(continued)

```

TGAGGGAATA TTTTGCCATC TGGGACTGTT GGTGCTGGG GGCAGTGGCT ATGAGCTCAG
  1990      2000      2010      2020      2030      2040

TTAATAAACT CAAGCAGTTT CCTTCCAAAC ACACATGTCC TACTTAACGT GTCCAACAGA
  2050      2060      2070      2080      2090      2100

-----
GATGATCATA CTCATANGCT GCTAAAACAT TANTTTTATT TTGAGAAAAG TCTATTTCATG
  2110      2120      2130      2140      2150      2160
-----
----- Alu insert -----
TTCTTGGCCC ATGGAGTTTT CATTINATTA NTTTATTTAT TTGTCAGAGA TGGAGTCTCA
  2170      2180      2190      2200      2210      2220

-----
CTATGTTGCT CAAGCTGGTC TCCAACCTCT GGGCTCAAGC GATCTTCCTA CTTTGGCCTT
  2230      2240      2250      2260      2270      2280

-----
TGAAAGCGCT GAGATTGCCT GTGTGAGCCA TCATGGGGGC TCACTGGCCC ACTGATTAAT
  2290      2300      2310      2320      2330      2340

CAGATTAATT GTTTTTTGCT ATTGAANTTG TTTGACTTCC TTGTATATTC GGATATTTAC
  2350      2360      2370      2380      2390      2400

CCATTCTAAC ACGTAGGGTT TGCAAATATT TTCTCTCATG TTCTGTGTG CCTTTTCACT
  2410      2420      2430      2440      2450      2460

CAGTTGATGG TTTCCTTTGC TGTGCAGGTG CTTTAGTGTT CAACGCAGCC CCGCTTGTCT
  2470      2480      2490      2500      2510      2520

ATTTTCCATT TTATTGCCTG TCCCTTTGAT GTCATAGCCA AGAAATAATT GCCCAGATTA
  2530      2540      2550      2560      2570      2580

ATGTCAAAAA GCTTTATCCC TATATATTCT TCTAGTAGTT TATGGTTTCA GATCTTATGT
  2590      2600      2610      2620      2630      2640

TTAGGTCTTC AATCCATTGA GTTGATTTTT GTATGTGGTA TAAGAAAAAA GACCACATGT
  2650      2660      2670      2680      2690      2700

ATACATATCT CAAATTCTAA GGTAGTATAT ATTAGACACA TACAATGTGT CTATTTACAC
  2710      2720      2730      2740      2750      2760

ACATTGAGCT GAAAATAATA AACATATTTT TATCTTTCAA TCAACTCTAT CTCTATCTCA
  2770      2780      2790      2800      2810      2820

CTGAACCTGT TTCACCTATA GCCTGATGAG GTTGCTGTCC TCTCTACCCC AGCTCCTATA
  2830      2840      2850      2860      2870      2880

GGAGACTGCT CATCCCCTAA CCTCAAAAAC CCCTTCATGA GGGTGATAAT GCCCTTGAAT
  2890      2900      2910      2920      2930      2940

```

Figure 1(continued)

CCTGCAATGA ATTAGTTCTC TACTACAGTG GAATTCAGGT CTGTTATGAG GGTCTGGATC
 2950 2960 2970 2980 2990 3000
 TCTGAAGAGA AGAGCTCTCA TTTTCAGAAA ATAAGCAGGA TTTATTCCCT GAAATTACTG
 3010 3020 3030 3040 3050 3060
 AATTAAATCA CTGTTTCGAT TACTTTTTGC AATATTAAAA GTAAATATTT AAACAGGTAA
 3070 3080 3090 3100 3110 3120
 AAACAGAAAT AATGGTAGGG TCCTTATCAT CACCGTGAAT TCCAAGCTAG CATAGACACT
 3130 3140 3150 3160 3170 3180
 AAACCTAGAG ATTCACTA GAATGAAAGC TGGGAGAGCA GAGGAGTCTC AGAAGGATGT
 3190 3200 3210 3220 3230 3240
 GGAGGCCAAT GGACACCTGC AACCTCTCCA ACGAAATGCC TACCTCCTCT CACTGCAGCA
 3250 3260 3270 3280 3290 3300
 TCCATCTCTG AGCCTTCTCG CAGCAGAGCT ATAAATTCAG CCTGGCTCCT CCGTTCCAC
 3310 3320 3330 3340 3350 3360
 Spsite CDS start
 ###<-----***-----
 ACATCCACTC CTGCTCTCCC TCCTCTCCTC CAGGTGACTA CAGTTATGAG GACCCTCACC
 3370 3380 3390 3400 3410 3420
 ----- Exon 2 -----
 CTCCTCTCTG CCTTTCTCCT GGTGGCCCTT CAGGCCTGGG CAGAGCCGCT CCAGGCAAGA
 3430 3440 3450 3460 3470 3480

 GCTCATGAGA TGCCAGCCCA GAAGCAGCCT CCAGCAGATG ACCAGGATGT GGTCAATTAC
 3490 3500 3510 3520 3530 3540
 Spsite
 ----->### ###
 TTTTCAGGAG ATGACAGCTG CTCTCTTCAG GTTCCAGGTG AGAGATGCCA GCATGCAGAG
 3550 3560 3570 3580 3590 3600
 CTACAGACTA GACAGAAGGA CAGGAGACAG GCTCTGGAAT TGGATCTCAG TGGCAGATGT
 3610 3620 3630 3640 3650 3660
 CACTTAGGTG GCTATACTTA ACATCTCTGG TCCTGGATTT TCTCATATCT AAATGGAATA
 3670 3680 3690 3700 3710 3720
 GAGAACCAAA GAAATCTAAG AGATTTTCTT TTCTCCAAA ACTTGATTCC AAGATATGAC
 3730 3740 3750 3760 3770 3780
 TGTGAAATTC ACTAGATTTA AGATATAAGG AGATGCTACC TAGTTCCTTC TGGAGCCAGA
 3790 3800 3810 3820 3830 3840

Figure 1(continued)

CAAACAAGCT TAAGTATATA GGAAAATATT TCACCCTGTC TATATAGGAG GTTTTAGAAC
 3850 3860 3870 3880 3890 3900
 CTGGAGAGGA GCCTAAGAAT GTGTTTCAGGT GTGTGTGTGA TGGGCAGGAA TGCAGAAAAG
 3910 3920 3930 3940 3950 3960
 TGAAGCAAAG GAGAATGAGT CTCGAATCCT GTGTGACCAG CACTGCTCTG TGTATTTATT
 3970 3980 3990 4000 4010 4020
 CCTATTGACT GAGATTGTTT GTGCTACCGG CTGTAATACA GCCAACATCA CTCATCAGCC
 4030 4040 4050 4060 4070 4080
 AACATGTGAC TTCTCCAAGA TTCCCTTTAC CACCCACTGC TGNACCCCGT ACTCAGTTTC
 4090 4100 4110 4120 4130 4140

Spsite
 ###<-----
 TGATGCTCTC TCTGGGTCCC CAGGCTCAAC AAAGGGCTTG ATCTGCCATT GCAGAGTACT
 4150 4160 4170 4180 4190 4200

----- Exon 3 -----
 ATACTGCATT TTTGGAGAAC ATCTTGGTGG GACCTGCTTC ATCCTTGGTG AACGCTACCC
 4210 4220 4230 4240 4250 4260

CDS stop
 ***-----
 AATCTGCTGC TACTAAGCTT GCAGACTAGA GAAAAAGAGT TCATAATTTT CTTTGAGCAT
 4270 4280 4290 4300 4310 4320

Poly Ad

 ----->
 TAAAGGGAAT TGTTATTCTT ATACCTTGTC CTCGATTTC TGTCTCATC CCAAATAAAT
 4330 4340 4350 4360 4370 4380
 ACTTGGTAAC ATGATTCCG GGTTTTTTTT TTTT
 4390 4400 4410

Figure 1(continued)

Figure 2

40045130 401304

```

      540      550      560      570      580      590
DEF4 AACAGCGACCACCCCTAGCGAGGCTTGAGATGCCTCTTCCCTCCTTAAGACTGAGAGCGC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX AACAGCGACCACCCCTAGCGAGGCTTGAGATGCCTCTTGCCTCCTTAAGACTGAGGGAGA
      540      550      560      570      580      590

      600      610      620      630
DEF4 CGCT-----GCCCCAGTCCCTCCATAGCCCAGTGCCTGGCTGCCTTCA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX CGCTTCAGCTCTCACTCCACTGCCCCAAGTCCTCCACAGCGCGGTGCCTG-CTGCCTTCA
      600      610      620      630      640      650

      640      650      660      670      680      690
DEF4 GCCAGAGCTGCAGGGG-AGGCCCTGAGCACCCAAGTCCTGCTGGACCAGCGCTGTGCACG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX CACAGAGCTGCAGGGGNAGGTCCTGTGTATCC--GGCCTGCTGGACCAGCGCTGTGCACA
      660      670      680      690      700      710

      700      710      720      730      740      750
DEF4 GCCCTCCCATGGCGGCAGGGGCTGCCTGGACTGCATACTGGGTTCAGCAACCTCACTATA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX ACCCTCCCATGGCAACAGTGGCTGCCCCGGCCTGCACACTGGGCTTGGCAACCTCGCTGTA
      720      730      740      750      760      770

      760      770      780      790      800      810
DEF4 GGTATTCAATCCCTCAGGAACAACCTGCATTCTTTCTCATTTCAGAAACCTCATCCCCT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX GGTATTTATTCCCTCAGGAGTGACTGCATTCTTTCCCATTTCCAGAAAACCTGATGCCAT
      780      790      800      810      820      830

      820      830      840      850      860
DEF4 TTACCTCACTACAAGGAGGAGGATG-----GTGGAGAGTGGTACATTTTAAAATGT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX TTACCTCACTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTGGTACATTTTAAAATGT
      840      850      860      870      880      890

      870      880      890      900      910      920
DEF4 GCACTAGTCTCCCTGGGACTCCCCTTCAAATAACCCAGGAGGGACCACACAAGGGAAAGC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX GCACTATTCTCCCTAGGACTCCCCTCAAATAACCCAGGAGGGACCATAACAGCTCATTC
      900      910      920      930      940      950

      930      940      950      960      970      980
DEF4 TTATGCATCCCCCCCACCC-AGTGACCATCTTCTAACTCTGGGTGTAGGGAGACTCGTA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX CTGTGTATCCCAAGCATANGAGTAATCATCCCACTCATGCTGAGTGTATGGTGGCCATTA
      960      970      980      990      1000      1010

      990      1000      1010      1020      1030      1040
DEF4 AGCCTACG--GGATTGGTTTGGGAACAGGGTATTTGAGCTCACAACACAAGGTGATGCAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX AGCCTGCCCTGAACTGGCTTTAGAACAAGGTGTTGAGCACACAGCACCG-----
      1020      1030      1040      1050      1060

```

Figure 2(continued)

Figure 2 (continued)

10045120.101201

```

      1580      1590      1600      1610      1620      1630
DEF4  TGATTGCTGCCTGGGCCTGGCCAGGGCTGACCAAAGTAGACGAGGGGTCCGTACCTCCGT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  TGGTAGCTACCTGGGCCTGGCCAGGGCTGACCA--TAGACAAGGCATC---CCTCTGT
      1590      1600      1610      1620      1630

      1640      1650      1660      1670      1680      1690
DEF4  GGACTCCTGCTTGAAGTCCAGCTTTCTGCCAAATTTCTCAACTGCCCTTGTTAACAGTTA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  GAACTCCTATTTTAATGCCAGCTTCCCAACAAATTTCTCAACTGCTCTTACCAGCAGGTA
      1640      1650      1660      1670      1680      1690

      CAAT box
      1700 ----1710      1720      1730      1740      1750
DEF4  TTAAAGTACCCAATAGAAAAGTAACGCTGAAAAATTAGGACACCTGATACCAAAAGACCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  TTAAACTACTCAATAGAAAAGTAACCTGAAAA-TTAGGACACCTGTTCCCAAAAGACCC
      1700      1710      1720      1730      1740      1750

      TATA box
      -----      1770      1780      1790      1800
DEF4  TTAAATAAGG-AAGTCCTCTC-CTCTGTGTGCATGGCTGCTCTTG---CTACATAAGACC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  TTAAATAGGGGAAGTCCTTTCNCTGCTTGTGCACAGCTGCTGATGTGGCAACATGAGGCC
      1760      1770      1780      1790      1800      1810

      mRNA start -->      SpSite
      1810      1820 | 1830      1840      1850      1860 ----
DEF4  TGGAACACAGGACTGTGTCTGCCCTCTCTGCTCGCCCTGCCTAGCTTGAGGATCTGTAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  TGGGACAGGGGACTGTCCTCTGCCCCTCTGGTAGCCTCACGTAGCTTAACAATCTGTCA
      1820      1830      1840      1850      1860      1870

      --      1880      1890      1900      1910      1920
DEF4  GTAACACAA-----AACTTAACTTTCACATTGAGGTTTCAATATTGAAGCTGTGTCCCC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  GTAATACAATACAAAACCTTAACTTTCATACTGCGGTTCCACCCAGGAAGCTGTGTTCCC
      1880      1890      1900      1910      1920      1930

      1930      1940      1950      1960      1970      1980
DEF4  AGTCTGACCTCTCACTGTGGGGCCACCCAGAGGACCCAGCGTGAAGCCCCTGCTGTGAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  AATCTGACCCGTGATTATGGGGCCACCTCAGAGGNAACCCAGTGAGGGAA-TATTTTG--
      1940      1950      1960      1970      1980      1990

      1990      2000      2010      2020      2030      2040
DEF4  CTTCTATCTGGGTGTCTGGCGGCTGCTGGGGGTAATGGCTACTAGCTAAGTCAATAGAGA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  ---CCATCTGGGA--CTGTTGGTTGCTGGGGGCAGTGGCTATGAGCTCAGTTAATA----
      2000      2010      2020      2030      2040

```

Figure 2 (continued)

Figure 2 (continued)

Figure 2 (continued)

Figure 2 (continued)

```

----->
      3350      3360      3370      3380      3390      3400
DEF4  CTCTTCAGGTTTTCAGGTGAGAGAGGCCAGCATAAAAAAGCTACCGAGTCTAGAGAGACGG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  CTCTTCAGGTTCCAGGTGAGAGATGCCAGCATGCAGA-GCTAC--AGACTAGACAGAAGG
      3570      3580      3590      3600      3610

      3410      3420      3430      3440      3450      3460
DEF4  ATGGGAGATGGGCTCTGGAATCACATCTCAATGGTGGATGTCACTTAGGTGGCTTTACTT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  ACAGGAGACAGGCTCTGGAATTGGATCTCAGTGGCAGATGTCACTTAGGTGGCTATACTT
      3620      3630      3640      3650      3660      3670

      3470      3480      3490      3500      3510      3520
DEF4  ACCATCTCTGGGCCTCGATTTTCTTATCTCGAACTGAATAGAGAGACAAACAAATGTAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  AACATCTCTGGTCTCGATTTTCTCATATCTAAATGGAATAGAGAACCAAGAAATCTAA
      3680      3690      3700      3710      3720      3730

      3530      3540      3550      3560      3570      3580
DEF4  GT-AGTCTTCTTTCTCCAAAGACTTGATTCCAAGGTATGTCTATAAAATTCGCTAGGGTT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  GAGATTTTCTTTCTCCAAAAGCTTGATTCCAAGATATGACTGTGAAATTCAGTAGATTT
      3740      3750      3760      3770      3780      3790

      3590      3600      3610      3620      3630
DEF4  AAGATATGGAGAGACAGATTGACCAGTTCTTTCTGGATCTAAACAAGTA-GAT--ATTAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  AAGATATAAGGAGATG--CTACCTAGTTCCTTCTGGAGCCAGACAAACAAGCTTAAGTAT
      3800      3810      3820      3830      3840      3850

      3640      3650      3660      3670      3680      3690
DEF4  AG-GGAAAATATTTCATTCTGCCAACAAAGGAAATTTTAAAAACTGGAGATGGGCTTAAG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  ATAGGAAAATATTTCACCCTGTCTATATAGGAGGTTTTAGAACCTGGAGAGGAGCCTAAG
      3860      3870      3880      3890      3900      3910

      3700      3710      3720      3730      3740      3750
DEF4  AGTATGTTTCAGGTGTGTGTCTGATGGGGCA--AAAGCACACAAATCAGAGCAAAAGAGAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  AATGTGTTTCAGGTGTGTGTGTGATGGG-CAGGAATGCAGAAAAGTGA-AGCAAAGGAGAA
      3920      3930      3940      3950      3960      3970

      3760      3770      3780      3790      3800      3810
DEF4  TGAGTCTCAAATCCTGTATGAGCAGCATTGCTCTGTGTATTTATTCCTATTGACTAAGGT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  TGAGTCTCGAATCCTGTGTGACCAGCACTGCTCTGTGTATTTATTCCTATTGACTGAGAT
      3980      3990      4000      4010      4020      4030

      3820      3830      3840      3850      3860      3870
DEF4  TGTGTGTGCTACCGGCACTAATGCAGCCAGCATCACCGGTCAGCCAGCATGTGCATTCTC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX  TGTGTGTGCTACCGGCTGTAATACAGCCAACATCACTCATCAGCCAACATGTGACTTCTC
      4040      4050      4060      4070      4080      4090

```

Figure 2 (continued)

10045130.101301

```

3880      3890      3900      3910      3920      3930
DEF4      CAAGATTCCCTTTACCACCCACCGCTGACCTTGGTGCTTAATTTCTCAGTCTTCCTCTGT
          :::::::::::::::::::::::::::: : : : : : ::::: : : :::::
DEFX      CAAGATTCCCTTTACCACCCACTGCTGNACCCCGTACTCAGTTTCTGATGCTCTCTCTGG
          4100      4110      4120      4130      4140      4150

          <----- exon3 ----->
3940      3950      3960      3970      3980      3990
DEF4      GTTCCCAGGCTCAACAAGGGGCATGGTCTGCTCTTGCGAGATTAGTATTCTGCCGCGAAC
          :::::::::::::::::::::::::::: : : : : : ::::: : : :::::
DEFX      GTCCCCAGGCTCAACAAGGGCTTGATCTGCCATTGCAGAGTACTATACTGCATTTTGG
          4160      4170      4180      4190      4200      4210

          ----- exon3 -----
4000      4010      4020      4030      4040      4050
DEF4      AGAAGTTCGTGTTGGGAAGTGCCTCATTGGTGGTGTGAGTTTCACATACTGCTGCACGCG
          ::::: : : : : ::::: : : : : : : : : : : : : : : : : : : : :
DEFX      AGAACATCTTGGTGGGACCTGCTTCATCCTTGGTGAACGCTACCCAATCTGCTGCT----
          4220      4230      4240      4250      4260      4270

          ----- exon3 -----
4060      4070      4080      4090      4100      4110
DEF4      TGTCGATTAACATTCTGCTGTCCAAGAGAATGTCATGCTGGGAACGCCATCATCGGTGGT
          : : :
DEFX      -----ACTAA-----

          ----- exon3 -----
4120      4130      4140      4150      4160      4170
DEF4      GTTAGCTTCACATGCTTCTGCAGCTGAGCTTGCGAGAATAGAGAAAAATGAGCTCATAATT
          :::::::::::::::::::: : : : : : :::::
DEFX      -----GCTTGCAGACTAGAGAAAAA-GAGTTCATAATT
          4280      4290      4300

          ----- exon3 -----
4180      4190      4200      4210      4220      4230
DEF4      TGCTTTGAGAGCTACAGGAAATGGTTGTTTCTCCTATACTTTGTCCTTAACATCTT-TCT
          : : : : : : : : : : : : : : : : : : : : : : : : : : :
DEFX      TTCTTTGAGCATTAAAGGGAATTGTTATT---CTTATACCTTGTCCTCGATTTCCTGTCC
          4310      4320      4330      4340      4350      4360

          Poly Ad
          ----->
4240      4250      4260      4270      4280      4290
DEF4      TGATCCTAAATATATATCTCGTAACAAGATGTCTTTGTTTACACCTCTTTGAAATTTGAT
          : : : : : : : : : : : : : : : : : : : : :
DEFX      TCATCCCAAATAAATACTTGGTAACATGATTTCGGGTTTTTTTTTTTTT
          4370      4380      4390      4400      4410

```

Figure 2(continued)

Figure 3

```

<----- Signal peptide ----->
      5      10      15      20
MetArgThrLeuThr LeuLeuSerAlaPhe LeuLeuValAlaLeu GlnAlaTrpAlaGlu

----- Propiece -----
      25      30      35      40
ProLeuGlnAlaArg AlaHisGluMetPro AlaGlnLysGlnPro ProAlaAspAspGln

----- Propiece -----
      45      50      55      60
AspValValIleTyr PheSerGlyAspAsp SerCysSerLeuGln ValProGlySerThr

-----><----- Mature peptide -----
      65      70      75      80
LysGlyLeuIleCys HisCysArgValLeu TyrCysIlePheGly GluHisLeuGlyGly

----- Mature peptide ----->
      85      90      94
ThrCysPheIleLeu GlyGluArgTyrPro IleCysCysTyr

```

Figure 4

	SIGNAL	PROPIECE
DEF4_HUMAN	MRIIALLAAILLVALQVRA	GPLQAR-----GDEAPGQ-EQRGPEQDISISFAWDKSS
DEF5_HUMAN	MRTIAILAAILLVALQAQA	ESLQER-----ADEATTQ-KQSGEDNQDLAISFAGNGLS
DEF6_HUMAN	MRTLTLTAVLLVALQAQA	EPLQAEDDPLQAKAYEADAQ-EQRGANDQDFAVSFAEDASS
DEF1_HUMAN	MRTLAILAAILLVALQAQA	EPLQAR-----ADEVAAPEQIAADIPEVVVSLAWDESL
DEFX	MRTLTLLSAFLVALQAWA	EPLQAR-----AHEMPAQ-KQPPADDQDVVIYFSGDDSC
	** ...*.* ***** *	** * *

	PROPIECE	Mature PEPTIDE
DEF4_HUMAN	ALQVSGSTRGM	VCSCRLVFCRRTLRVGNCLIGGVSTFYCCTTRVD
DEF5_HUMAN	ALRTSGSQARA	TCYCRTGRCATRESLSGVCEISGRLYRLCCR---
DEF6_HUMAN	SLRALGSTRAF	TCHCRR-SCYSTEYSYGTCTVMGINHRFCCL---
DEF1_HUMAN	APKHPGSRKNM	ACYCRIPACIAGERRYGTCTIYQGRWLWAFCC---
DEFX	SLQVPGSTKGL	ICHCRVLYCIFGEHLGGTCFILGERYPICCY---
	. . **	* ** * * * *
		^ ^ ^ ^ ^ ^ ^

Figure 5